National Data Warehouse Area Preparatory Guide

Customer's Guide and Reference Manual to Prepare for the Initial Data Load to the National Data Warehouse

Version 1.0

June 2004

Department of Health and Human Services

Indian Health Service

Information Technology Support Center (ITSC)

TABLE OF CONTENTS

Introduction

Welcome applicable facility/area staff to the Area Preparatory Guide. This guide outlines the steps needed to prepare for sending data for inclusion in the IHS National Data Warehouse (NDW) repository. This guide contains all the currently known information to support this project from each facility. It will evolve, be revised and new versions will be posted on the NDW website. These changes may be software patches, various updates, new helpful information, or processes improved by actual experience. We have processed one area at this writing.

Getting Help with the Guide

Please send any questions or suggestions regarding this formatting guide to:

E-Mail: itscdatawarehouse@ihs.gov
Website: http://www.ndw.ihs.gov/

Sending Data to NDW

Once in full production, the process will FTP the files to the NDW automatically in the background. However, because this process is currently in beta, the files will need to FTP'd manually. The instructions for the address and login will be supplied to the Area ISC.

Getting Started

Three weeks before your scheduled start date, the Area ISC needs to provide to the NDW the contact name, facility name, ASUFAC, phone number, and email address. We need this data to supply: an email acknowledgement to each facility that we have received their file(s); another email notice when their data is loaded into the NDW and to view the website for a summary of data verifications or discrepancies, and support issues or file rejections where we may need to call the contact.

Begin surveying all area facilities for what RPMS software updates are needed and apply those updates. Please see the following section titled "RPMS Software Requirements". In assigning a schedule of when/who to run processes, organize facilities by level of difficulty and have those with the most hurdles start earlier in the Area's cycle.

The NDW will provide to the Area staff a schedule of when to start each facility's exports. Your input into creating this schedule is very important.

If the number of visits per month and the total number of registration records is known by facility, they can derive how long the processing might take and how much space is required in MB to perform the initial exports. See the section titled "Runtime and HD Space Statistics" to compute your facilities needs.

The initial load, or "back load," to the NDW consists of a one time full registration extract and visit extract(s) containing visits from October 1, 2000 to the present. The NDW extract is created using HL7 messages. HL7 messages are fairly large and will use up disk space during the initial load process. These HL7 messages will be dumped to a host file for export to the NDW computer. The facility will need to manage the archiving of the files that will be created. Once the NDW acknowledges loading the data that they received, the file can be deleted.

The visit extract takes up space in the Cache environment (^INTHU global and temp globals) and creates an output file that needs to be ftp'd to the NDW computer. The INTHU global will be purged every few days so the space needed is temporary and will be freed up once the purger is run. The

purger is set to purge all HL7 messages from the INTHU global once they are 3 days old. Once the output file is FTP'd to the NDW computer, it can be zipped, archived or deleted.

The basic process for running initial loads (vs. incremental) is the following: (If you need more information or detailed instructions on running the processes, see the Technical User Guide)

1. Install the IHS/RPMS NDW software package using kids build "bdw_0100.k"

2. Install GIS patch 12 using kids build "gis_0301.12k", "gis_0301.12n", and "gis_0301.12g"

3. Assign BDW security keys to the users responsible for NDW initial loads and for incremental NDW exports. The keys are:
    a. BDWZMENU (unlocks main menu)
    b. BDWZ RESET (allows re-setting a log entry)
    c. BDWZ REG EXPORT (allows user to do a full registration export)

4. Run the full Registration initial load using option ARG Data Warehouse Full Registration Export. This option can be found on BDWMENU using the patch "BLM-ARG". Two files will be created and placed into the site's defined export directory. The first file will be called "BDWDWPXasufac.julian date" and the second file will be titled "asufacdatetime.BDW". An example of these files from SX are:
    a. "0001113040530112045.BDW" and "BDWDWPX000111.123"
5. Perform the visit initial load. The visit extract is created using the BLV menu option. Each facility will need to determine, based on the size of the site and available disk space, how many visits to process with each run. See section titled "Runtime and HD Space Statistics".

RPMS Software Requirements

The following patch updates should be followed in this order.

1. GIS Version 3.01
2. GIS Version 3.01 patch 2
3. Fileman version 22
4. Kernel 8.0 or later
5. IHS Patient Dictionaries (AUPN) v 99.1 through patch 13 (when patch 14 is released that will be a requirement)
6. IHS VA Support Files (AVA) v 93.2 through patch 18. If a PIMS user, patch 18 must be immediately followed by an install of PIMS, if not PIMS user yet, patch 18 must be installed.
7. PCC Management Reports v3.0 through patch 15
8. RPMS Registration v7.0 through patch 2

Runtime and Hard Disk Space Statistics

The following statistics were gathered using the first three facilities data which is a very limited history. Other factors that could skew the results would be the type of hardware, processor and throughput speeds, amount of memory or HD space, and the number of applications running on the same server or at the same time. Our data assumes the exports were executed after hours and no other processes were running. It may also be skewed because most of the visit data is San Xavier's and they have a fast machine. Your results will most likely be slower than these.

1. Registration - Initial Loads
   This process creates an audit and an extract file and will be created once. To derive a runtime you will need to know approximately how many registration records you have at your facility.

a) Runtime
Where Y = the number of these records:
(Y * 0.000610869) = the total runtime to create both files.
Example: If you have 200,000 registration records, the formula is:
(200,000*.000610869) = 122 minutes. 122 minutes divided by 60 minutes = 2.04 hours.

b) Space in MB
Using the number of records again as Y. (Y * 0.001031234) = Total space in MB of both files in total.
Example: If you have 70,000 registration records, the formula is:
(70,000*.001031234) = 72.19MB (megabytes of disk space).

2. Encounters (Visits) - Initial Loads
This process creates an audit and an extract file in the same procedure. It differs from registrations in that there can be many more records so the processing times will be greater. Because of this fact, the process can be run by a date range to keep it manageable. Unless you're one of the biggest facilities, a safe rule to follow is to run six months the first time to see how that works. To derive a runtime or find MB space, you will need to know approximately how many visits you have at your facility each month on average.

a) Runtime
Where Y = the number of these records each Month * 6 (for six months):
(Y * 0.001115622) = the total runtime to create both files.
Example: If you have 7,000 visits each month, the formula is:
(42,000*0.001115622) = 46.86 minutes

b) Space in MB
Using the number of records times the six months again as Y. (Y * 0.001196255) = Total space in MB of both files in total.
Example: If you have 7,000 visits each month, the formula is
(42,000* 0.001196255) = 50.24MB (megabytes of disk space).

Area NDW Checklist

Create Area contact list... ☐

Survey of RPMS software update status.. ☐

Received a schedule from the NDW staff... ☐

Received the FTP instructions from the NDW staff... ☐

Performed calculation by facility for runtime and space constraints...................... ☐

Run the process for creating the registration extract audit files .... ☐

FTP these registration files to the NDW.. ☐

Run the process for creating the visit extract and audit files................................. ☐

FTP these visit files to the NDW.. ☐